



#8

OIPE

RAW SEQUENCE LISTING

DATE: 09/20/2002

PATENT APPLICATION: US/09/918,485

TIME: 11:00:18

Input Set : N:\Crf3\RULE60\09918485.raw

Output Set: N:\CRF4\09202002\I918485.raw

SEQUENCE LISTING

ENTERED

3 (1) GENERAL INFORMATION:

(i) APPLICANT: SANCHIS, Vincent
 LERECLUS, Didier
 MENU, Ghislaine
 LECADET, Marguerite-Marie
 MARTOURET, Daniel
 DEDONDER, Raymond

(ii) TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING FOR
 POLYPEPTIDES ENDOWED WITH A LARVICIDAL ACTIVITY TOWARDS
 LEPIDOPTERA

16 (iii) NUMBER OF SEQUENCES: 2

18 (iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
 (B) STREET: P.O. Box 1404
 (C) CITY: Alexandria
 (D) STATE: Virginia
 (E) COUNTRY: USA
 (F) ZIP: 22313

26 (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
 (B) COMPUTER: IBM PC compatible
 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

32 (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US/09/918,485
 (B) FILING DATE: 25-Oct-2001
 (C) CLASSIFICATION:

54 (vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US/08/461,551
 (B) FILING DATE: 05-JUN-1995
 (A) APPLICATION NUMBER: US 08/251,652
 (B) FILING DATE: 31-MAY-1994
 (A) APPLICATION NUMBER: US 07/458,754
 (B) FILING DATE: 11-DEC-1989
 (A) APPLICATION NUMBER: EP 88 401 121.4
 (B) FILING DATE: 06-MAY-1988
 (A) APPLICATION NUMBER: FR 87 08090
 (B) FILING DATE: 10-JUN-1987

58 (viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: HUNTINGTON, R. D.
 (B) REGISTRATION NUMBER: 27,903
 (C) REFERENCE/DOCKET NUMBER: 010830-073

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/918,485

DATE: 09/20/2002

TIME: 11:00:18

Input Set : N:\Crif3\RULE60\09918485.raw

Output Set: N:\CRF4\09202002\I918485.raw

```

63      (ix) TELECOMMUNICATION INFORMATION:
64          (A) TELEPHONE: (703) 836-6620
65          (B) TELEFAX: (703) 836-2021
68 (2) INFORMATION FOR SEQ ID NO: 1:
70      (i) SEQUENCE CHARACTERISTICS:
71          (A) LENGTH: 2711 base pairs
72          (B) TYPE: nucleic acid
73          (C) STRANDEDNESS: single
74          (D) TOPOLOGY: linear
76      (ii) MOLECULE TYPE: DNA (genomic)
80      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
82 AAGCTTCAAT AGAATCTCAA ATCTCGATGA CTGCTTAGTC TTTTAAATAC TGTCTACTTG      60
84 ACAGGGGTAG GAACATAATC GGTCAATTTT AAATATGGGG CATATATTGA TATTTTATAA      120
86 AATTTGTTAC GTTTTTTGTA TTTTTCATA AGATGTGTCA TATGTATTAA ATCGTGGTAA      180
88 TGAAAAACAG TATCAAAC TAAGAACTTT GGTAGTTTAA TAAAAAACG GAGGTATTTT      240
90 ATGGAGGAAA ATAATCAAAA TCAATGCATA CCTTACAATT GTTTAAGTAA TCCTGAAGAA      300
92 GTACTTTTGG ATGGAGAACG GATATCAACT GGTAATTACT CAATTGATAT TTCTCTGTCA      360
94 CTTGTTTCACT TTCTGGTATC TAACTTTGTA CCAGGGGGAG GATTTTCTAGT TGGATTAATA      420
96 GATTTTGTAT GGGGAATAGT TGGCCCTTCT CAATGGGATG CATTCTAGT ACAAATTGAA      480
98 CAATTAATTA ATGAAAGAAT AGCTGAATTT GCTAGGAATG CTGCTATTGC TAATTTAGAA      540
100 GGATTAGGAA ACAATTTCAA TATATATGTG GAAGCATTTA AAGAATGGGA AGAAGATCCTT      600
102 AATAATCCAG CAACCAGGAC CAGAGTAATT GATCGCTTTC GTATACTTGA TGGGCTACTT      660
104 GAAAGGGACA TTCCTTCGTT TCGAATTTCT GGATTGGAAG TACCCCTTTT ATCCGTTTAT      720
106 GCTCAAGCGG CCAATCTGCA TCTAGCTATA TTAAGAGATT CTGTAATTTT TGGAGAAAGA      780
108 TTGGGATTGA CAACGATAAA TGTCAATGAA AACTATAATA GACTAATTAG GCATATTGAT      840
110 GAATATGCTG ATCACTGTGC AAATACGTAT AATCGGGGAT TAAATAATTT ACCGAAATCT      900
112 ACGTATCAAG ATTGATAAAC ATATAATCGA TTACGGAGAG ACTTAACATT GACTGTATTA      960
114 GATATCGCCG CTTTCTTTCC AAACATGAC AATAGGAGAT ATCCAATTCA GCCAGTTGGT      1020
116 CAACTAACAA GGGAAGTTTA TACGGACCCA TTAATTAATT TTAATCCACA GTTACAGTCT      1080
118 GTAGCTCAAT TACCTACTTT TAACGTTATG GAGAGCAGCG CAATTAGAAA TCCTCATTTA      1140
120 TTTGATATAT TGAATAATCT TACAATCTTT ACGGATTGGT TTAGTGTTGG ACGCAATTTT      1200
122 TATTGGGGAG GACATCGAGT AATATCTAGC CTTATAGGAG GTGGTAACAT AACATCTCCT      1260
124 ATATATGGAA GAGAGGCGAA CCAGGAGCCT CCAAGATCCT TTACTTTTAA TGGACCGGTA      1320
126 TTTAGGACTT TATCAATTCC TACTTTACGA TTATTACAGC AACCTTGCCA GCGCCACCAT      1380
128 TTTAATTTAC GTGGTGGTGA AGGAGTAGAA TTTTCTACAC CTACAAATAG CTTTACGTAT      1440
130 GCAGGAAGAG GTACGGTTGA TTCTTTAACT GAATTACCGC CTGAGGATAA TAGTGTGCCA      1500
132 CCTCGCGAAG GATATAGTCA TCGTTTATGT CATGCAACTT TTGTTCAAAG ATCTGGAACA      1560
134 CCTTTTTTAA CAACTGGTGT AGTATTTTCT TGGACGCATC GTAGTGCAAC TCTTACAAAT      1620
136 ACAATTGATC CAGAGAGAAT TAATCAAATA CCTTTAGTGA AAGGATTTAG AGTTTGGGGG      1680
138 GGCACCTCTG TCATTACAGG ACCAGGATTT ACAGGAGGGG ATATCCTTCG AAGAAATACC      1740
140 TTTGGTGATT TTGTATCTCT ACAAGTCAAT ATTAATTCAC CAATTACCCA AAGATACCGT      1800
142 TTAAGATTTT GTTACGCTTC CAGTAGGGAT GCAGCAGTTA TAGTATTAAC AGGAGCGGCA      1860
144 TCCACAGGAG TGGGAGGCCA AGTTAGTGTA GATATGCCTC TTCAGAAAAC TATGGAAATA      1920
146 GGGGAGAACT TAACATCTAG AACATTTAGA TATACCGATT TTAGTAATCC TTTTTCATTT      1980
148 AGAGCTAATC CAGATATAAT TGGGATAAGT GAACAACCTC TATTTGGTGC AGGTTCTATT      2040
150 AGTAGCGTTG AACTTTATAT AGATAAAATT GAAATTATTC TAGCAGATGC AACATTTGAA      2100
152 GCAGAATCTG ATTTAGAAAG AGCACAAAAG GCGGTGAATG CCCTGTTTAC TTCTTCCAAT      2160
154 CAAATCGGGT TAAAAACCGA TGTGACGGAT TATCATATTG ATCAAGTATC CAATTTAGTG      2220
156 GATTGTTTAT CAGATGAATT TTGTCTGGAT GAAAAGCGAG AATTGTCCGA GAAAGTCAAA      2280

```

RAW SEQUENCE LISTING

DATE: 09/20/2002

PATENT APPLICATION: US/09/918,485

TIME: 11:00:18

Input Set : N:\Crif3\RULE60\09918485.raw

Output Set: N:\CRF4\09202002\I918485.raw

```

158 CATGCGAAGC GACTCAGTGA TGAGCGGAAT TTA CTTC AAG ATCCAAACTT CAGAGGGATC 2340
160 AATAGACAAC CAGACCGTGG CTGGAGAGGA AGTACAGATA TTACCATCCA AGGAGGAGAT 2400
162 GACGTATTCA AAGAGAATTA CGTCACACTA CCGGGTACCG TTGATGAGTG CTATCCAACG 2460
164 TATTTATATC AGAAAATAGA TGAGTCGAAA TTA AAAAGCTT ATACCCGTTA TGAATTAAGA 2520
166 GGGTATATCG AAGATAGTCA AGACTTAGAA ATCTATTTGA TCGCGTACAA TGCAAAACAC 2580
168 GAAATAGTAA ATGTGCCAGG CACGGGTTCC TTATGGCCGC TTTCAGCCCA AAGTCCAATC 2640
170 GGAAAGTGTG GAGAACCGAA TCGATGCGCG CCACACCTTG AATGGAATCC TGATCTAGAT 2700
172 TGTTCTGCA G 2711
174 (2) INFORMATION FOR SEQ ID NO: 2:
176 (i) SEQUENCE CHARACTERISTICS:
177 (A) LENGTH: 823 amino acids
178 (B) TYPE: amino acid
179 (C) STRANDEDNESS: unknown
180 (D) TOPOLOGY: unknown
182 (ii) MOLECULE TYPE: peptide
186 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
188 Met Glu Glu Asn Asn Gln Asn Gln Cys Ile Pro Tyr Asn Cys Leu Ser
189 1 5 10 15
192 Asn Pro Glu Glu Val Leu Leu Asp Gly Glu Arg Ile Ser Thr Gly Asn
193 20 25 30
196 Ser Ser Ile Asp Ile Ser Leu Ser Leu Val Gln Phe Leu Val Ser Asn
197 35 40 45
200 Phe Val Pro Gly Gly Gly Phe Leu Val Gly Leu Ile Asp Phe Val Trp
201 50 55 60
204 Gly Ile Val Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile Glu
205 65 70 75 80
208 Gln Leu Ile Asn Glu Arg Ile Ala Glu Phe Ala Arg Asn Ala Ala Ile
209 85 90 95
212 Ala Asn Leu Glu Gly Leu Gly Asn Asn Phe Asn Ile Tyr Val Glu Ala
213 100 105 110
216 Phe Lys Glu Trp Glu Glu Asp Pro Asn Asn Pro Ala Thr Arg Thr Arg
217 115 120 125
220 Val Ile Asp Arg Phe Arg Ile Leu Asp Gly Leu Leu Glu Arg Asp Ile
221 130 135 140
224 Pro Ser Phe Arg Ile Ser Gly Phe Glu Val Pro Leu Leu Ser Val Tyr
225 145 150 155 160
227 Ala Gln Ala Ala Asn Leu His Leu Ala Ile Leu Arg Asp Ser Val Ile
228 165 170 175
231 Phe Gly Glu Arg Trp Gly Leu Thr Thr Ile Asn Val Asn Glu Asn Tyr
232 180 185 190
235 Asn Arg Leu Ile Arg His Ile Asp Glu Tyr Ala Asp His Cys Ala Asn
236 195 200 205
239 Thr Tyr Asn Arg Gly Leu Asn Asn Leu Pro Lys Ser Thr Tyr Gln Asp
240 210 215 220
243 Trp Ile Thr Tyr Asn Arg Leu Arg Arg Asp Leu Thr Leu Thr Val Leu
244 225 230 235 240
246 Asp Ile Ala Ala Phe Phe Pro Asn Tyr Asp Asn Arg Arg Tyr Pro Ile
247 245 250 255
250 Gln Pro Val Gly Gln Leu Thr Arg Glu Val Tyr Thr Asp Pro Leu Ile

```

RAW SEQUENCE LISTING

DATE: 09/20/2002

PATENT APPLICATION: US/09/918,485

TIME: 11:00:18

Input Set : N:\Crif3\RULE60\09918485.raw

Output Set: N:\CRF4\09202002\I918485.raw

| | | | | | | |
|-----|---|---|-----|-----|-----|-----|
| 251 | | 260 | | 265 | | 270 |
| 254 | Asn Phe Asn | Pro Gln Leu Gln Ser Val Ala Gln Leu Pro Thr Phe Asn | | | | |
| 255 | | 275 | | 280 | | 285 |
| 258 | Val Met Glu Ser Ser Ala Ile Arg Asn Pro His Leu Phe Asp Ile Leu | | | | | |
| 259 | | 290 | | 295 | | 300 |
| 262 | Asn Asn Leu Thr Ile Phe Thr Asp Trp Phe Ser Val Gly Arg Asn Phe | | | | | |
| 263 | 305 | | 310 | | 315 | 320 |
| 265 | Tyr Trp Gly Gly His Arg Val Ile Ser Ser Leu Ile Gly Gly Gly Asn | | | | | |
| 266 | | 325 | | 330 | | 335 |
| 269 | Ile Thr Ser Pro Ile Tyr Gly Arg Glu Ala Asn Gln Glu Pro Pro Arg | | | | | |
| 270 | | 340 | | 345 | | 350 |
| 273 | Ser Phe Thr Phe Asn Gly Pro Val Phe Arg Thr Leu Ser Ile Pro Thr | | | | | |
| 274 | | 355 | | 360 | | 365 |
| 277 | Leu Arg Leu Leu Gln Gln Pro Cys Gln Arg His His Phe Asn Leu Arg | | | | | |
| 278 | | 370 | | 375 | | 380 |
| 281 | Gly Gly Glu Gly Val Glu Phe Ser Thr Pro Thr Asn Ser Phe Thr Tyr | | | | | |
| 282 | 385 | | 390 | | 395 | 400 |
| 284 | Arg Gly Arg Gly Thr Val Asp Ser Leu Thr Glu Leu Pro Pro Glu Asp | | | | | |
| 285 | | 405 | | 410 | | 415 |
| 288 | Asn Ser Val Pro Pro Arg Glu Gly Tyr Ser His Arg Leu Cys His Ala | | | | | |
| 289 | | 420 | | 425 | | 430 |
| 292 | Thr Phe Val Gln Arg Ser Gly Thr Pro Phe Leu Thr Thr Gly Val Val | | | | | |
| 293 | | 435 | | 440 | | 445 |
| 296 | Phe Ser Trp Thr His Arg Ser Ala Thr Leu Thr Asn Thr Ile Asp Pro | | | | | |
| 297 | | 450 | | 455 | | 460 |
| 300 | Glu Arg Ile Asn Gln Ile Pro Leu Val Lys Gly Phe Arg Val Trp Gly | | | | | |
| 301 | 465 | | 470 | | 475 | 480 |
| 303 | Gly Thr Ser Val Ile Thr Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu | | | | | |
| 304 | | 485 | | 490 | | 495 |
| 307 | Arg Arg Asn Thr Phe Gly Asp Phe Val Ser Leu Gln Val Asn Ile Asn | | | | | |
| 308 | | 500 | | 505 | | 510 |
| 311 | Ser Pro Ile Thr Gln Arg Tyr Arg Leu Arg Phe Arg Tyr Ala Ser Ser | | | | | |
| 312 | | 515 | | 520 | | 525 |
| 315 | Arg Asp Ala Arg Val Ile Val Leu Thr Gly Ala Ala Ser Thr Gly Val | | | | | |
| 316 | | 530 | | 535 | | 540 |
| 319 | Gly Gly Gln Val Ser Val Asn Met Pro Leu Gln Lys Thr Met Glu Ile | | | | | |
| 320 | 545 | | 550 | | 555 | 560 |
| 322 | Gly Glu Asn Leu Thr Ser Arg Thr Phe Arg Tyr Thr Asp Phe Ser Asn | | | | | |
| 323 | | 565 | | 570 | | 575 |
| 326 | Pro Phe Ser Phe Arg Ala Asn Pro Asp Ile Ile Gly Ile Ser Glu Gln | | | | | |
| 327 | | 580 | | 585 | | 590 |
| 330 | Pro Leu Phe Gly Ala Gly Ser Ile Ser Ser Gly Glu Leu Tyr Ile Asp | | | | | |
| 331 | | 595 | | 600 | | 605 |
| 334 | Lys Ile Glu Ile Ile Leu Ala Asp Ala Thr Phe Glu Ala Glu Ser Asp | | | | | |
| 335 | | 610 | | 615 | | 620 |
| 338 | Leu Glu Arg Ala Gln Lys Ala Val Asn Ala Leu Phe Thr Ser Ser Asn | | | | | |
| 339 | | 625 | | 630 | | 635 |
| 341 | Gln Ile Gly Leu Lys Thr Asp Val Thr Asp Tyr His Ile Asp Gln Val | | | | | |
| 342 | | 645 | | 650 | | 655 |

RAW SEQUENCE LISTING

DATE: 09/20/2002

PATENT APPLICATION: US/09/918,485

TIME: 11:00:18 .

Input Set : N:\Crif3\RULE60\09918485.raw

Output Set: N:\CRF4\09202002\I918485.raw

```

345   Ser Asn Leu Val Asp Cys Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys
346               660               665               670
349   Arg Glu Leu Ser Glu Lys Val Lys His Ala Lys Arg Leu Ser Asp Glu
350               675               680               685
353   Arg Asn Leu Leu Gln Asp Pro Asn Phe Arg Gly Ile Asn Arg Gln Pro
354               690               695               700
357   Asp Arg Gly Trp Arg Gly Ser Thr Asp Ile Thr Ile Gln Gly Gly Asp
358   705               710               715               720
360   Asp Val Phe Lys Glu Asn Tyr Val Thr Leu Pro Gly Thr Val Asp Glu
361               725               730               735
364   Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys Ile Asp Glu Ser Lys Leu Lys
365               740               745               750
368   Ala Tyr Thr Arg Tyr Glu Leu Arg Gly Tyr Ile Glu Asp Ser Gln Asp
369               755               760               765
372   Leu Glu Ile Tyr Leu Ile Ala Tyr Asn Ala Lys His Glu Ile Val Asn
373               770               775               780
376   Val Pro Gly Thr Gly Ser Leu Trp Pro Leu Ser Ala Gln Ser Pro Ile
377   785               790               795               800
379   Gly Lys Cys Gly Glu Pro Asn Arg Cys Ala Pro His Leu Glu Trp Asn
380               805               810               815
383   Pro Asp Leu Asp Cys Ser Cys
384               820

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/918,485

DATE: 09/20/2002

TIME: 11:00:19

Input Set : N:\Crf3\RULE60\09918485.raw

Output Set: N:\CRF4\09202002\I918485.raw

L:33 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:34 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]